

Учреждение образования “Белорусский государственный экономический университет”

УТВЕРЖДАЮ

Ректор учреждения образования
“Белорусский государственный
экономический университет”

В.Н.Шимов

“ 30 ” 7.06. 2017 г.

Регистрационный № УД 5325-19 /уч.

КОМПЬЮТЕРНАЯ ЛИНГВИСТИКА

Учебная программа учреждения высшего образования по учебной дисциплине для специальности 1-25 81 10 «Экономическая информатика»

СОСТАВИТЕЛИ:

Могиленских Н.П., заведующий кафедрой межкультурной экономической коммуникации учреждения образования «Белорусский государственный экономический университет», кандидат филологических наук, доцент;

Кацнельсон И.А., старший преподаватель кафедры межкультурной экономической коммуникации учреждения образования «Белорусский государственный экономический университет».

РЕЦЕНЗЕНТЫ:

Новик Н.А., заведующий кафедрой профессионально ориентированной английской речи учреждения образования «Белорусский государственный экономический университет», кандидат филологических наук, доцент;

Тупик Д.И., декан факультета межкультурных коммуникаций учреждения образования «Минский государственный лингвистический университет», кандидат филологических наук, доцент.

РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:

Кафедрой межкультурной экономической коммуникации учреждения образования «Белорусский государственный экономический университет»

(протокол №11 от 24.05.2017 г.);

Научно-методическим советом учреждения образования «Белорусский государственный экономический университет»

(протокол № 5 от 21.06. 2017 г.).

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

«Компьютерная лингвистика» предназначена для обучения студентов I курса специальности 1-25 81 10 «Экономическая информатика» в рамках магистерской программы «Цифровые бизнес-коммуникации».

Компьютерная лингвистика является сегодня одной из наиболее активно развивающихся областей научных исследований и прикладных разработок. Это связано с растущим влиянием Интернета и появлением огромного количества новых технических устройств, важнейшей частью которых являются естественно-языковые интерфейсы. Это многоязычный информационный поиск, машинный перевод, автоматический анализ содержания документов, извлечение знаний, распознавание речи и т.п.

Цель и задачи учебной дисциплины

Целью освоения учебной дисциплины «Компьютерная лингвистика» является формирование профессиональной компетенции через ознакомление студентов с одним из новейших лингвистических направлений, его истоками, особенностями формирования и перспективами дальнейшего развития, а также овладение необходимыми навыками в области применения достижений компьютерной лингвистики для решения лингвистических учебно-познавательных задач, а также приемами их внедрения в практическую деятельность переводчика.

Основные задачи учебной дисциплины

- сформировать профессиональные компетенции в области лингвистических основ информатики и аспектах моделирования языка и мышления в компьютерной среде с помощью компьютерных программ;
- сформировать систему знаний, умений и навыков в области использования компьютерных технологий как в обучении языку, так и в научно-исследовательской деятельности будущих магистров;
- обучить магистрантов использованию и применению средств ИКТ в профессиональной деятельности специалиста, работающего в системе образования;
- ознакомить с современными приемами и методами использования средств ИТ при проведении разных видов учебных занятий, реализуемых в учебной и внеучебной деятельности;
- формировать умения работать с информацией, развитие коммуникативных способностей;

- подготовить будущего магистра к жизни в «информационном обществе», что предполагает формирование исследовательских умений, навыков применения оптимальных решений.

Требования к освоению учебной дисциплины

Результатом освоения учебной дисциплины «Компьютерная лингвистика» должна стать система знаний, обеспечивающих понимание сложного многоканального процесса речевой деятельности, опирающейся на взаимосвязь подсистем иерархически организованной системы языка, а также формирование компетенций, обеспечивающих применение полученных знаний в практической деятельности в области автоматической обработки лингвистической информации.

В результате обучения студенты должны

ЗНАТЬ:

- историю компьютерной лингвистики;
- место компьютерной лингвистики среди других лингвистических дисциплин;
- методы исследования компьютерной лингвистики;
- свойства сочетаемости лингвистических знаков различных уровней и особенности функционирования знаковых систем;
- типы структурной организации текстов;
- основные принципы и стратегии машинной обработки лингвистической информации, основные принципы построения искусственных языков и машинных словарей, разработки информационных банков, построения алгоритмов обработки текстов. предмет, объект и основные понятия синтактики, понятийный и терминологический аппарат изучаемой дисциплины;
- основные понятия и этапы автоматизированной обработки лингвистической информации.

УМЕТЬ:

- осуществлять сбор, создание, хранение, систематизацию, анализ, распространение, интерпретацию текстовой информации в системе «человек – компьютер – человек».

- анализировать структурно-синтаксическую систему изучаемых языков (АК - 2);
- проводить структурно-функциональный анализ текстов различной природы (АК - 2);
- классифицировать языки по морфологическим признакам (АК - 2);
- использовать методы синтактики в практических целях (АК - 3);
- выявлять синтаксические связи между единицами в языке и речи (АК - 3);
- идентифицировать явления языка и речи, которые служат для структурирования высказывания (АК - 3);
- систематизировать данные корпусной лингвистики проводить математико-статистическую обработку лингвистической информации (АК - 3);
- работать самостоятельно (АК - 4).

ВЛАДЕТЬ:

- методикой машинного автоматизированного перевода, навыками формального компьютерного моделирования естественного языка (АК - 1);
- правилами создания текстов различной функциональной направленности на изучаемых языках в соответствии с коммуникативной задачей (АК - 1);
- приемами и методами формально-синтаксического анализа предложения на родном и иностранном языке (АК - 2);
- навыками, связанными с использованием технических устройств, управлением информацией и работой с компьютером (АК - 7);
- навыками работы со специальной литературой, анализировать и оценивать собранную информацию, формировать информационно-аналитические базы данных (ПК - 15);
- приемами мозгового штурма, кейс-стади и игровыми технологиями (ПК - 25).

Межпредметные связи

Учебная дисциплина строится на принципах координации межпредметных связей со смежной учебной дисциплиной «Системы искусственного интеллекта».

Изучение учебной дисциплины «Компьютерная лингвистика» проводится на запланированных учебным планом аудиторных занятиях.

Всего часов по учебной дисциплине – 66, из них аудиторных часов – 30, в том числе 8 часов – лекции, 22 часов – практические занятия.

Форма контроля – зачет.

СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

В пределах учебной дисциплины «Компьютерная лингвистика» выделяются следующие разделы, темы и вопросы:

I Компьютерная лингвистика как научная дисциплина

Тема 1. Задачи компьютерной лингвистики в изучении естественного языка (ЕЯ)

История научного направления. Классификация направлений лингвистики. Направления современных исследований. Исследовательские группы, институты, конференции.

Тема 2. Формальные методы классификации полнотекстовых документов

Математическая постановка задачи распознавания образов и классификации. Формальные методы определения сходства ЕЯ документов на различных уровнях лингвистического анализа (морфологическом, синтаксическом, семантическом): кластерный анализ, деревья принятия решений, векторные методы, Байесовский классификатор. Применение методов классификации для задач определения авторства текстов.

II Естественные языки и их описание

Тема 3. Естественные и искусственные коммуникативные системы

Естественные и искусственные языки, их взаимодействие и использование. Метод лингвистического моделирования. Язык и метаязык. Методы структурной лингвистики, кибернетики и информатики (понятие вторичной моделирующей системы). Количественная и структурная стороны информации. Удельный вес синтаксических закономерностей в формализованных языках математической логики и математики.

Тема 4. Структурное описание естественных языков разных типов

Способы представления синтаксической структуры, их достоинства и недостатки. Дистрибутивный анализ. Дерево зависимостей. Дерево непосредственно составляющих. Трансформационный синтаксис. Генеративная грамматика и генеративная семантика. Структурирование грамматической семантики. Глубинные и поверхностные структуры. Синтактика частей речи. Части речи и классы слов.

III Синтаксический аспект текста

Тема 5. Синтактика текста

Предмет лингвистики текста. Единицы текста: сложное синтаксическое целое (ССЦ), сверхфразовое единство (СФЕ), абзац. Основные категории: когезия (когерентность), континуум, перспекция/ретроспекция, завершенность. Функционально-смысловые типы текстов: описание, повествование, рассуждение. Структура связанного текста, средства связности. Структурно-семантические модели различных типов текста. Типология текстов.

Тема 6. Лингвистический процессор. Алгоритмы лингвистического разбора и анализа текста. Парсеры ЕЯ-предложений

Лингвистический процессор - функциональная структура. Методы морфологического анализа, используемые в лингвистических процессорах. Морфологические словари. Алгоритмы синтаксического и семантического анализа для автоматических систем обработки текстов. Парсеры ЕЯ. Прикладные системы - спэлчекеры, текстовые редакторы, системы профессионального редактирования.

Тема 7. Автоматическая обработка текста

Общие представления об автоматизированной обработке речи (устной и письменной). Понятие об алгоритме. Лингвистический алгоритм. Алгоритмизация обработки языковой информации. Ввод языковой информации в компьютер. Системы распознавания устной речи. Автоматический морфологический анализ. Словарь основ и словарь словоформ. Автоматический синтаксический анализ словосочетания и предложения. Проблемы генерации предложений. Автоматическое индексирование, классифицирование, аннотирование и реферирование текстов.

Тема 8. Лингвистические проблемы создания систем искусственного интеллекта

Представление знаний в интеллектуальных системах. Системы понимания естественного языка. Современный машинный перевод. Принципы построения и использования переводческих машинных словарей. Машинный фонд национального языка и его основные компоненты. Терминологические банки данных. Способы формального представления знаний: логические правила вывода, семантические сети, фреймы. Тенденции развития систем искусственного интеллекта.

**УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ «КОМПЬЮТЕРНАЯ ЛИНГВИСТИКА»
(ДНЕВНОЕ ОТДЕЛЕНИЕ)**

Номер раздела, темы	Название раздела, темы	Количество аудиторных часов					Иное	Форма контроля знаний
		Лекции	Практические занятия	Семинарские Занятия	Лабораторные Занятия	Количество часов УСР		
1	2	3	4	5	6	7	8	9
I	Компьютерная лингвистика как научная дисциплина	4	6					
1	Задачи компьютерной лингвистики в изучении естественного языка (ЕЯ)	2	2				[1; 4]	Устный опрос Тест
2	Формальные методы классификации полнотекстовых документов	2	4				[2; 4]	Устный опрос Презентация
II	Естественные языки и их описание	2	4					
3	Естественные и искусственные		2				[3; 5]	Портфолио

	коммуникативные системы							Тест
4	Структурное описание естественных языков разных типов	2	2				[2; 5]	Устный опрос Тест
III	Синтаксический аспект текста	2	12					
5	Синтактика текста		2				[1; 4]	Устный опрос Презентация
6	Лингвистический процессор. Алгоритмы лингвистического разбора и анализа текста. Парсеры ЕЯ-предложений	2	4				[2; 3]	Устный опрос Презентация
7	Автоматическая обработка текста		2				[3; 5]	Устный опрос Проект Тест
8	Лингвистические проблемы создания систем искусственного интеллекта		4				[3; 4]	Устный опрос Проект Тест
	Всего часов	8		22				зачет

**УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ «КОМПЬЮТЕРНАЯ ЛИНГВИСТИКА»
(ЗАОЧНОЕ ОТДЕЛЕНИЕ)**

Номер раздела, темы	Название раздела, темы	Количество аудиторных часов					Иное	Форма контроля знаний
		Лекции	Практические занятия	Семинарские Занятия	Лабораторные Занятия	Количество часов УСР		
1	2	3	4	5	6	7	8	9
I	Компьютерная лингвистика как научная дисциплина	2	2					
1	Задачи компьютерной лингвистики в изучении естественного языка (ЕЯ)	2					[1; 4]	Устный опрос Тест
2	Формальные методы классификации полнотекстовых документов		2				[2; 4]	Устный опрос Презентация
II	Естественные языки и их описание		4					
3	Естественные и искусственные		2				[3; 5]	Портфолио

	коммуникативные системы							Тест
4	Структурное описание естественных языков разных типов		2				[2; 5]	Устный опрос Тест
III	Синтаксический аспект текста	2	6					
5	Синтактика текста		2				[1; 4]	Устный опрос Презентация
6	Лингвистический процессор. Алгоритмы лингвистического разбора и анализа текста. Парсеры ЕЯ-предложений		2				[2; 3]	Устный опрос Презентация
7	Автоматическая обработка текста		2				[3; 5]	Устный опрос Проект Тест
8	Лингвистические проблемы создания систем искусственного интеллекта	2					[3; 4]	Устный опрос Проект Тест
	Всего часов	4		12				зачет

ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

Методические рекомендации по организации самостоятельной работы студентов по учебной дисциплине «Компьютерная лингвистика»

В овладении знаниями учебной дисциплины важным этапом является самостоятельная работа студентов. Рекомендуется бюджет времени для самостоятельной работы в среднем 2-2,5 часа на 2-х часовое аудиторное занятие.

Основными направлениями самостоятельной работы студента являются:

- первоначально подробное ознакомление с программой учебной дисциплины;
- ознакомление со списком рекомендуемой литературы по дисциплине в целом и ее разделам, наличие ее в библиотеке и других доступных источниках, изучение необходимой литературы по теме, подбор дополнительной литературы;
- изучение и расширение лекционного материала преподавателя за счет специальной литературы, консультаций;
- подготовка к практическим занятиям по специально разработанным планам с изучением основной и дополнительной литературы;
- подготовка к выполнению диагностических форм контроля (тесты, коллоквиумы, контрольные работы и т.п.);
- подготовка к зачету.

ЛИТЕРАТУРА

Основная:

1. Гальперин, И.Р. Текст как объект лингвистического исследования / И.Р.Гальперин. - М., 2005. – 139 с.
2. Зубов, А.В. Основы искусственного интеллекта для лингвистов: учеб пособие / А.В.Зубов. – М.: Университетская книга; Логос, 2007. – 319 с.
3. Марчук, Ю.Н. Компьютерная лингвистика: учеб. пособие / Ю.Н.Марчук. – М.: АСТ: Восток-Запад, 2007. – 317 с.
4. Пиотровский, Р.Г. Синергетика текста: учеб. пособие. / Р.Г.Пиотровский. - Минск: МГЛУ, 2005. – 156 с.
5. Зубов, А.Б., Зубова, И.И. Информационные технологии в лингвистике / А.В.Зубов. - М.: «Академия», - 2014. – 167 с.

Дополнительная:

6. Баранов, А.Н. Введение в прикладную лингвистику / А.Н.Баранов. - М., 2012. – 212 с.
7. Беляев, М.И., Григорьев, С.Г. Теория и практика создания образовательных электронных изданий / М.И.Беляев // М.: Изд-во РУДН, – 2003. – 241 с.
8. Гриншкун, В.В. Григорьев, С.Г. Образовательные электронные издания и ресурсы / В.В.Гришкун // Учебно-методическое пособие для студентов педагогических вузов и слушателей системы повышения квалификации работников образования. - Курск: КГУ, М.: МГПУ – 2010. – 98 с.
9. Теньер, Л. Основы структурного синтаксиса: Пер. с фр. М.: Прогресс, 1988. – 653 с.
10. Бондарко, А.В. Функциональная грамматика. / А.В. Бондарко. – Л.: «Наука», 1984. – 136 с.
11. Вежбицкая, А. Семантические универсалии и описание языков. М.: «Языки русской культуры», 1999. – 780 с. // А. Вежбицкая. Семантические универсалии и описание языков. М.: «Языки русской культуры», 1999; Раздел I. Из книги «Семантика: примитивы и универсалии». – с. 3 – 88; Раздел II. Семантика грамматики. Глава «Лексические прототипы как универсальное основание межъязыковой идентификации «частей речи». – С . 134-170.
12. Норман, Б.Ю. Синтаксис речевой деятельности. / Б.Ю. Норман. – Минск: Вышэйшая школа, 1978. – 131 с.

ПРОТОКОЛ СОГЛАСОВАНИЯ УЧЕБНОЙ ПРОГРАММЫ УВО

Название учебной дисциплины, которой требуется согласование	Название кафедры	Предложения об изменениях в содержании учебной программы учреждения высшего образования по учебной дисциплине	Решение, принятое кафедрой, разработавшей учебную программу (с указанием даты и номера протокола)
Системы искусственного интеллекта	Кафедра информационных технологий	Нет <i>С.И.Ш.</i> <i>М. И. Сагдоева -</i> <i>Зав. каф. ИТ</i>	Протокол № 11 от 24 мая 2017 г

ДОПОЛНЕНИЯ И ИЗМЕНЕНИЯ К УЧЕБНОЙ ПРОГРАММЕ УВО
на ____ / ____ учебный год

№ п/п	Дополнения и изменения	Основание

Учебная программа пересмотрена и одобрена на заседании кафедры

_____ (протокол № ____ от _____ 201_ г.)
(название кафедры)

Заведующий кафедрой

_____ (ученая степень, ученое звание) _____ (подпись) (И.О.Фамилия)

УТВЕРЖДАЮ
Декан факультета

_____ (ученая степень, ученое звание) _____ (подпись) (И.О.Фамилия)